

---

# Konzeption und Aufbau der digitalen Bibliothek Information und Medien

---

Projektbericht des Seminars „Digitale Bibliothek“

Im Auftrag der Projektgruppe erarbeitet von

**Florian Engster** und **Markus Klimek**

unter Mitarbeit von

**Daniel Zimmer**

August 2001

*(veröffentlicht September 2001)*

Titelangaben für dieses Dokument:

**Engster, Florian:**

Konzeption und Aufbau der digitalen Bibliothek Information und Medien : Projektbericht des Seminars „Digitale Bibliothek“ / Florian Engster ; Markus Klimek ; Daniel Zimmer. – Stuttgart : Fachhochschule Stuttgart – Hochschule der Medien, 2001. – 27 S.

NE: Klimek, Markus;; Zimmer, Daniel:

Details zur Verfügbarkeit des hier vorgestellten Projektes finden sich in Abschnitt 3 auf Seite 23.

Als Online-Publikation unter <http://diana.iuk.hdm-stuttgart.de/digbib/> abrufbar.

Copyright © 2001 by Florian Engster <[engster@hbi-stuttgart.de](mailto:engster@hbi-stuttgart.de)>, Markus Klimek <[klimek@hbi-stuttgart.de](mailto:klimek@hbi-stuttgart.de)>, Daniel Zimmer <[zimmer@hbi-stuttgart.de](mailto:zimmer@hbi-stuttgart.de)>.

Dieser Text darf zu akademischen Zwecken frei kopiert werden. Eine weitere Nutzung über die Regelungen des Urheberrechts hinaus oder zu kommerziellen Zwecken ist untersagt.

#### **Kurzreferat**

Die „digitale Bibliothek Information und Medien“ stellt ein konzeptionelles und technisches System dar, um digitale Information zu speichern, aufzubereiten und zu vermitteln. Seine Umsetzung erfolgte im Rahmen eines Projektseminars an der ehemaligen FH Stuttgart – Hochschule für Bibliotheks- und Informationswesen (HBI). Die Herangehensweise und Einrichtung des Systems werden knapp als Projektbericht erläutert.

#### **Schlagwörter (SWD)**

s.Elektronische Bibliothek ; s.Planungskonzept ; s.Errichtung

#### **Schlagwörter (INFODATA)**

Bibliothek ; Informationsversorgung ; Digital ; Planung

#### **Notation (ACM CCS 98)**

H.3.7 ; H.1.1



# Inhaltsverzeichnis

<b>Einleitung</b>	<b>7</b>
<b>1 Rahmenbedingungen</b>	<b>9</b>
1.1 Warum digitale Bibliotheken? . . . . .	9
1.2 Das Konzept einer digitalen Bibliothek . . . . .	9
1.3 Probleme und Chancen . . . . .	11
1.4 Konzeption für das Projekt . . . . .	13
<b>2 Umsetzung</b>	<b>15</b>
2.1 Hyperwave Information Server . . . . .	15
2.2 Greenstone Digital Library Software . . . . .	16
2.3 Resumée . . . . .	20
<b>3 Ausblick</b>	<b>23</b>
<b>Literaturverzeichnis</b>	<b>25</b>
<b>Kolophon</b>	<b>27</b>



## Einleitung

Nicht zuletzt auf dem Bibliothekartag 2001 in Bielefeld wurde das Thema diskutiert, ob im Zeitalter des Internet der Typus einer Bibliothek eigentlich nicht schon überholt ist. Die Wissensgesellschaft der Gegenwart kann mit wenigen Mausklicks vom heimischen Arbeitsplatz aus auf das Wissen der Welt zugreifen, und dies ohne dass jemand gewissenhaft einen Katalog gepflegt hat und die Dokumente vor Ort sorgsam vorhält. Niemand ist mehr abhängig vom Gutwillen der Auskunftsbibliothekarin, von der Kenntnis der Kataloge, deren Aufbau gleichsam den Regeln eines geheimen Zirkels nur wenigen bekannt ist, unbelastet von den Öffnungszeiten eines altherwürdigen Gebäudes oder schlichtweg von der Tatsache, dass Bücher wie jedes organische Material der Verrottung anheim gegeben sind. Und der Segen der Suchmaschinen erlöst den informationssuchenden Nutzer vom Zwang, sich über formale Titel, Autorennamen und dergleichen den Kopf zu zerbrechen, da er jeden Satz, jedes Wort in Millionen von Web-Seiten der Welt finden kann. Ohne das Haus verlassen zu müssen.

In diese Irre gehen zumindest populäre Orakelien (etwa Zimmer 2000).

Dass sich die Zukunft der qualifizierten Informationsvermittlung nicht auf einen netzfähigen Rechner mit einem WWW-Client reduzieren lässt, war den Teilnehmern des Seminars „Digitale Bibliothek“ bereits bewusst. Ihr Ansporn war die Nutzung der Möglichkeiten, die sich durch eine weltweite Netzwerkinfrastruktur ergeben, ihre Herausforderung waren die Probleme, die ein mehrerer Terabytes großer, dezentral und unkontrolliert betriebener Datenbestand mit sich bringt.

In der Zeit der zwei Semester zwischen Oktober 2000 und Juli 2001 beschäftigte sich die Gruppe mit den Konzepten digitaler Bibliotheken und letztendlich mit dem Aufbau einer Modellanwendung.

Das Ergebnis war ein Projekt, dass sich in drei Phasen gliedert:

**Phase 1: Konzeption** Hier wurde zunächst das Konzept der digitalen Bibliothek definiert, die Rahmenbedingungen formuliert und eine Arbeitsgrundlage geschaffen.

**Phase 2: Aufbau** Der Aufbau umfasste die Bewertung verfügbarer Software-Lösungen, die Sammlung und Einstellung von geeignetem Material und den Aufbau einer Server-Anwendung zum Testbetrieb.

**Phase 3: Produktivbetrieb** Die letzte Phase des Projektes wäre der Übergang in den Produktiv- oder Echtbetrieb, in dem die digitale Bibliothek dann öffentlich nutzbar ist.

Ziel des Seminars war es, das Projekt bis Phase 2 zu betreuen und im zeitlich machbaren Rahmen der Semester zu dokumentieren. Auf diese Weise wäre ein solider und vorzeigbarer Dienst auf die Beine gestellt, dessen Nachnutzung dann im Ermessen der Bibliothek bzw. des Informationstechnischen Service läge.

Der vorliegende Bericht soll vorrangig *keine technische Dokumentation* des Projektes sein, sondern möchte in erster Linie die strategischen und konzeptionellen Überlegungen auf dem Weg zu einer funktionierenden digitalen Bibliothek darstellen. Bei einer technisch orientierten Anwendung wie einer digitalen Bibliothek können Schwerpunkte auf diesem Gebiet allerdings nicht ausbleiben.

Das Seminar unter der Leitung von Prof. Margarete Payer wurde getragen von Studierenden der Fächer „Wissenschaftliche Bibliotheken“ (WB) und „Informationsmanagement“ (IM), die auch einen Großteil der Arbeit selbstständig erledigten. Die schwerpunktmäßige Betreuung des hier vorgestellten Projektes lag bei:

- Florian Engster (WB)
- Annette Gehde (WB)
- Claudius Hörth (IM)
- Markus Klimek (IM)
- Helga Möck (WB)
- Birgit Patzelt (WB)
- Daniel Zimmer (WB)

Die Veröffentlichung dieses Berichtes fällt noch in die Aufbauphase des Projektes. Änderungen an den hier dargestellten Sachverhalten sind daher immer noch möglich, obgleich diese sich nur noch auf Details beziehen sollten.

Unser kollektiver Dank gilt an dieser Stelle allen Personen, die auf vielfältige Weise dieses Projekt unterstützt haben.

Stuttgart, im August 2001



# 1 Rahmenbedingungen

Bibliotheken dienen in erster Linie der Versorgung von Forschung und Lehre mit wissenschaftlich relevanter Information. Eine mit diesem Anspruch konzipierte digitale Bibliothek vermeidet, dass sie um ihrer Selbst willen aufgebaut wird oder mit strategischen Problemen ihrer Zweckbestimmung zu kämpfen hat.

Im Folgenden sind daher unsere Überlegungen angeführt, auf Grund derer wir uns in Abschnitt 2 der praktischen Umsetzung widmen.

## 1.1 Warum digitale Bibliotheken?

Die in den letzten Jahren aufgekommene Informationsflut manifestiert sich vorrangig im Internet. So war (nach Witten u. a. 1999) im Jahr 1993 täglich ein Volumen von 100 Megabytes im weltweiten Datennetz neu verfügbar. Eine Zahl, die etwa 400 gedruckten Büchern entspricht. 1999 war dieser tägliche Zuwachs bereits auf 1 Gigabyte gewachsen.

Die Masse der Daten allein ist jedoch nicht das Problem. Dieses stellt sich erst, wenn es um das Finden relevanter Information in diesen Unmengen an größtenteils unstrukturierten Daten geht.

Ein erster Versuch der Bewältigung war die Einrichtung von Suchdiensten, welche die Server des WWW absuchen und aus den Inhalten der dort vorliegenden Dateien Volltextindizes bilden. Zu einer Zeit, als das Internet und das WWW vorrangig akademisch genutzt wurden, war diese „Erschließung“ der Datenmengen noch brauchbar. Heutzutage verwässern nicht-relevante Dokumente<sup>1</sup>, die immerhin den Großteil des WWW-Inhaltes darstellen, die Ergebnisse, so dass in vielen Fällen die Suche nach relevanter Information dem Nutzer anstrengende und zeitraubende Recherchetaktiken abverlangt, die ihn vielfach überfordern.

Unser Anspruch an digitale Bibliotheken besteht darin, die Informationsflut des Internet in fachlichem Kontext zu bewältigen und für digitale Informationsressourcen das zu leisten, was eine konventionelle Bibliothek für konventionelle Informationsressourcen leistet. Mit ihnen soll ein durchdachter Ansatz geliefert werden, Suchende schnell und zuverlässig zu relevanter Information zu führen. Dabei steht das Motto „Klasse statt Masse“ über allem Bemühen.

## 1.2 Das Konzept einer digitalen Bibliothek

Doch wie werden digitale Bibliotheken diesen Anforderungen, hier einmal nur auf die konzeptionelle Struktur bezogen, gerecht?

Um uns eine Metapher zu erlauben: wenn man den Buch- und Zeitschriftenmarkt auf das Internet abbildet, eine konventionelle Bibliothek als riesige Volltextdatenbank und ihre Kataloge als die Indizes dazu versteht, was der Sache durchaus angemessen entspricht, so gelangt man recht einfach zu einer möglichen funktionalen Vorstellung einer digitalen Bibliothek, die wie folgt aussehen könnte:

**Erschließung digitaler Ressourcen:** Information muss zwei Eigenschaften erfüllen: sie muss verwertbar und nachvollziehbar auffindbar sein. Letzte-

---

<sup>1</sup>Wir gebrauchen den Begriff „Dokument“ synonym für jede hier mögliche Art einer Medieneinheit.

res bedeutet auch, dass sie zitierbar sein muss. Durch einen eindeutigen Nachweis einer Ressource, also durch ihre formale Erschließung, ist diese eindeutig gekennzeichnet, beschrieben und kommunizierbar. Durch ihre inhaltliche Erschließung kann eine schnelle Aussage über ihre Inhalte getroffen werden, die thematische Gültigkeit und dadurch über ihre Verwertbarkeit für einen bestimmten Zweck. Beides sind Verfahren, die sich bei der Erschließung konventioneller Materialien schon lange bewährt haben.

**Archivierung digitaler Ressourcen:** keine Bibliothek weist Bestände nach, die sie selbst nicht vor Ort hat. Und die Schnelllebigkeit des Internet, sowie die problemlose Veränderbarkeit digitalen Materials bringt noch das Problem mit sich, dass Änderungen häufig nicht wahrgenommen oder nachvollzogen werden können. Die lokale Speicherung digitaler Ressourcen archiviert diese in ihrem Zustand zu einem bestimmten, festzuhaltenden Zeitpunkt und gibt so gegebenenfalls diese Ressource in verschiedenen Versionen wieder. Der Vergleich mit der Handhabung verschiedener Ausgaben konventioneller Materialien verdeutlicht diesen Vorteil. Zudem werden die eingestellten Ressourcen im Rahmen der technischen Möglichkeiten und Gegebenheiten dauerhaft archiviert, und sind auch nach einem Verschwinden vom Originalort noch zugreifbar.

**Angebot fachlich relevanter Inhalte:** Massenbewältigung für sich genommen ist nur ein technisches und infrastrukturelles Problem. Kompetenz ist in Bereichen der inhaltlichen Bewertung und qualitativen Auswahl gefragt. Diesem Anspruch müssen digitale Bibliotheken ebenso wie ihre konventionellen Vorbilder genügen. Im starren Gebilde konventionellen wissenschaftlichen Publizierens gab es einfache Anhaltspunkte, die Relevanz von Quellen zu bewerten. Das anarchistische und pluralistische Prinzip des Internet ermöglicht es allerdings, relevante Information außerhalb dieser Wege zu veröffentlichen, mit dem Nebeneffekt, dass es kaum formale Kriterien gibt, diese von irrelevanter oder falscher Information zu unterscheiden. Dieser vielleicht heikelste Punkt verlangt Sachkompetenz und klar definierte, gläserne Qualitätsfaktoren, um für den Nutzer eine verlässliche Größe darzustellen, mit der entsprechenden Sicherheit im Umgang mit den angebotenen Ressourcen.

Zur Schaffung dieser Sicherheit gehört die zusammen mit dem Einstellen der Dokumente in die Sammlung der digitalen Bibliothek einhergehende Klärung des Copyright und der Nutzungsrechte. Mit der Schaffung des Zugangs zu ihrem Angebot stellt die digitale Bibliothek die Verwendbarkeit der so verfügbaren Materialien im Rahmen der entsprechenden rechtlichen Bedingungen sicher.

**Digitale Bibliotheken sind zusätzliche Angebote:** wir erachten den Typus einer digitalen Bibliothek nur als tragbar, wenn er von einer bestehenden Einrichtung mit entsprechender fachlicher Kompetenz umgesetzt wird. Konkret bedeutet dies auch, dass digitale Bibliotheken eines von vielen Angeboten einer konventionellen Bibliothek sein sollten. Dies ermöglicht ein integrierendes Angebot an verschiedenen Formen von Information, und die Fähigkeiten der Mitarbeiter können in diesem Bereich fachverwandt eingesetzt werden.

## 1.3 Probleme und Chancen

Ebenso wenig, wie nicht alles Glänzende Gold ist, gibt es die perfekten Lösungen für Probleme. Bei der Umsetzung der oben angeführten Ansprüche stößt man immer wieder auf Schwierigkeiten und Probleme, die häufig durch die digitale Natur der Materialien begründet sind. Daneben jedoch zeichnen sich auch durchaus wertvolle Chancen ab.

### 1.3.1 Formatfragen

Das Internet wurde als systemunabhängiges weltweites Datennetz entworfen und Tim Berners-Lee entwickelte das WWW vor dem Hintergrund, den Austausch von Daten über die am CERN existierenden Systemgrenzen hinweg zu ermöglichen (s. Berners-Lee 1999). Die inzwischen jedoch neu gezogenen Grenzen betreffen die Austauschbarkeit von Dokumentformaten. Eine aktuelle Referenz zu den existierenden Dateiformaten schlägt mit 1127 Seiten zu Buche (s. Born 2000).

Ein Vorteil der Offenheit des Internet ist die zwangsläufig eintretende Konzentration auf offene Formate. Als Quasi-Standards zur elektronischen Publikation von textuellen Dokumenten haben sich die Formate HTML<sup>2</sup>, PDF, PostScript und Microsoft Word herausgebildet. Ferner sind dies  $\text{\TeX}$ / $\text{\LaTeX}$  und SGML/XML-Anwendungen. Bis auf das Word-Format sind alle Formate dokumentiert (W3C 1995 u. ö.; Adobe 2001, 2000; DANTE 1994 u. ö.; ISO 1986; W3C 2000) und frei verwendbar. Entsprechend groß ist auch ihre Verbreitung und Unterstützung.

Problematisch hierbei ist der Umgang mit einer heterogenen Masse an digitalem Material. Oberstes Anliegen ist natürlich der Erhalt des Informationswertes, erst diesem nachgeordnet folgt der Erhalt des Originallayouts.

Formate, die ihre Inhalte strukturiert speichern sind daher von großem Wert, da bei diesem Ansatz die visuelle Formatierung zweitrangig ist. Dies ist bei SGML/XML-Anwendungen und  $\text{\TeX}$ / $\text{\LaTeX}$  der Fall. PDF, PostScript und Microsoft Word zielen jedoch auf eine layoutgetreue Darstellung ab. Schwierigkeiten ergeben sich erst bei Fragen der Umkopie in andere Formate zur besseren Darstellung auf anderen Geräten oder für eine von der Darstellung losgelösten Verarbeitung.

Für eine digitale Bibliothek sind die Formate HTML und PDF daher von besonderem Interesse, da sie inzwischen allgemein verbreitet und im Endeffekt gut verwertbar sind. Bei PDF besteht der Vorteil darin, auf verschiedenen Systemen das gleiche Layout zu präsentieren, und gleichzeitig recht einfach in reinen Text oder HTML überführbar zu sein. Methoden wie die Verschlüsselung von PDF-Dateien erschweren hier aber wieder den Umgang, da diese oft auch nicht ausgedruckt werden können und in der Regel kein Text exportiert werden kann. Die Volltextindizierung und weitere Verwertung solcher Dateien ist damit schlicht unmöglich.

Die Speicherung von Bildern und Grafiken bewegt sich im Bereich des WWW in fassbaren Maßen. Hier dominieren die Formate JPEG, PNG, GIF und SVG. Die bei der Verwendung GIF (Graphics Interchange Format) auftretenden Patentprobleme haben dessen Einsatz fraglich gemacht und waren der Anlass zur Entwicklung des patentfreien PNG Formates (Portable Network Graphics). Das

---

<sup>2</sup>HTML sei hier als spezielle SGML/XML-Anwendung gesondert genannt.

SVG Format (Scalable Vector Graphics) wurde vom W3C entwickelt und speichert XML-basiert Vektorgrafiken.

### 1.3.2 Beschaffbarkeit und Nachweis

Wir führten oben bereits das Problem der Auffindbarkeit digitaler Ressourcen im Internet an. Dieses Problem beschäftigt uns nun an dieser Stelle. Während digitale Bibliotheken Ordnung in das Chaos des Internet bringen sollen, und auf diese Weise den Nutzer bei der Informationsrecherche entlasten und unterstützen sollen, muss genau diesem Problem bei der Pflege des Bestandes begegnet werden.

Häufig findet sich relevantes Material weltweit verteilt auf Servern von Fachgesellschaften, anderen Bibliotheken, Preprint- und Verlagsservern oder im privaten Angebot von Fachpersonen, und Hinweise darauf in Newsgroups und E-Mail-Listen.

Die klassische Recherche unter Zuhilfenahme von Fachdatenbanken, Katalogen oder fachlichen als auch allgemeinen Suchdiensten darf hierbei auch nicht außen vor bleiben.

Eine Herausforderung für den bibliografischen Nachweis sind die besonderen Anforderungen an die Beschreibung digitaler Ressourcen. Mit Dublin Core (Dublin Core 1999) und einer dazu entsprechenden Norm (DIN e. V. u. a. 2001), sowie den RAK-NBM (DBI 1996 u. ö.) existieren zwar zum Teil weit verbreitete und vielfach unterstützte Ansätze zu einer einheitlichen Beschreibung solcher Materialien, jedoch weisen diese Ansätze teils erhebliche Mängel oder schlichtweg ungelöste Fragen auf, so dass ein sinnvoller Einsatz erschwert wird.

Im Bibliothekswesen wird immer noch mit dem Anspruch der Perfektion katalogisiert. In vielen Fällen ist dieser Anspruch auch sinnvoll und gerechtfertigt. Der diffuse und oben schon als schwierig dargestellte Bereich der digitalen Ressourcen verlangt von Erschließungs- und Nachweisregeln einerseits Flexibilität, andererseits durch die aus der Sache heraus gegebenen Internationalität auch ein verlässliches Niveau.

Beachtet werden muss, dass sowohl Personal als auch Kunden im Grunde genommen noch in Kategorien alter Regelwerke denken, und so lange die Bestände konventioneller Bibliotheken noch danach erfasst werden, ist es von Vorteil, konventionelle und digitale Materialien vergleichbar zu erschließen. Das Internet stellt ein Medium dar, das fluide und unberechenbar ist. Während ein Buch oder eine Zeitschrift noch mit der Hand fassbar ist, muss im Bereich der digitalen Ressourcen mit Objekten gearbeitet werden, deren Struktur körperlos ist. Ein Buch hat zum Beispiel eine feste, unveränderliche Sequenz an Seiten, die alle zusammen genommen dieses Buch bilden. Wie aber handhabt man mehrere zu einer Arbeit gehörenden HTML-Seiten? Als einzelne Seiten? Als einzelne Kapitel?

### 1.3.3 Datenvolumen

In diesem Bereich steckt die unschlagbare Stärke digitaler Bibliotheken. Diese können ohne Platzprobleme Riesenmengen an Daten speichern und dank immer leistungsfähigerer Technik auch immer schneller handhaben.

Eine Volltext-Datenbank aller bis zum Jahre 1975 weltweit erschienenen Bücher könnte auf einem Speichervolumen von 15 Terabytes, also 15 000 Giga-

byte, in komprimierter und indizierter Form untergebracht werden (nach Witten u. a. 1999). Nachdem bereits Festplatten mit 100 Gigabytes Volumen für Heimcomputersysteme allmählich üblich werden, würde ein Festplattencluster von 150 Platten das gesamte bis 1975 veröffentlichte Weltwissen fassen. Beim Einsatz von gewöhnlichen Heimcomputern mit jeweils 4 Festplatten, benötigte man in etwa 40 Systeme, die auf einer Fläche von 20 qm bequem unterzubringen wären.

Die Verwaltung eines solchen Datenvolumens stellt hinsichtlich der Leistungsfähigkeit moderner Computersysteme und des Einsatzes ausgeklügelter Techniken zur Indizierung und Komprimierung von Bild- und Textdaten immer weniger ein Problem dar. Witten u. a. (1999) haben dies anschaulich dargestellt.

#### 1.3.4 Qualität

Tim Berners-Lees Vision des WWW war es, „dass potentiell alles mit allem verknüpft ist.“ (Berners-Lee 1999, S. 11), und dass jeder eigene Inhalte einstellen kann, während Ted Nelsons Xanadu-Projekt (Nelson u. a. 1960 u. ö.) eher kontrollierend ausgelegt war. Leider – oder zum Glück – hat sich das WWW durchgesetzt.

Dies führt uns in ein Medium, dessen Inhalte von theoretisch jeder Person weltweit eingestellt werden können. Das traditionelle Publizieren im akademischen Bereich umfasste in den meisten Fällen ein Peer-Reviewing, setzte voraus, dass der Autor renommiert war, und der sogenannte „*Impact-Factor*“ bestimmte die Qualität einer Zeitschrift. Dies fehlt in hohem Maße im Bereich des WWW. Zwar sind diese formalen Kriterien noch greifbar, wenn es um die digitalen Angebote z. B. von Fachverlagen geht, doch aus der hohen Unzufriedenheit von Autoren und Nutzern wissenschaftlicher Publikationen entstand eine immer höhere Akzeptanz und beiderseitige Nutzung von Preprint- und Freeprint-Severn (vgl. Biersch 2000) und weniger moderierten Publikationskanälen.

Hinter diesem Bereich wird es schwierig, formale Kriterien zur Bewertung digitaler Ressourcen fest zu machen. Die Fähigkeit „alles mit allem“ zu verknüpfen macht es möglich, relevante und nicht-relevante Ressourcen in eine Sequenz zu schließen. Vielfach ist es nicht möglich, ohne eingehende Beschäftigung und Prüfung des Materials, eine Aussage über seine Qualität zu treffen. Für die Pflege einer digitalen Bibliothek wird bei der Einstellung von Ressourcen aus diesem „offenen“ Bereich eine hohe Sachkompetenz erforderlich, die mit Sorgfalt und kritischer Prüfung einhergehen muss.

### 1.4 Konzeption für das Projekt

Die hier angesprochenen Punkte lassen in ihrem Kern keine anderen Tätigkeiten als die klassischen bibliothekarischen und dokumentarischen erkennen. Vielmehr werden diese gerade im Umfeld der digitalen Bibliotheken um so mehr benötigt, oft in einer verfeinerten Form.

Bislang haben wir uns mit der konzeptionellen Seite digitaler Bibliotheken beschäftigt und betont, wie wichtig in diesem Zusammenhang bibliothekarische Arbeit ist. Wertvoll kann eine digitale Bibliothek, wie ihr konventionelles Vorbild, nur sein, wenn sie entsprechende Dienstleistungen und Inhalte anbietet.

Für unser Projekt mündete dies in einem Angebot, das auf den Informationsbedarf des hiesigen Fachbereiches ausgerichtet ist, also auf eine Situation,

die bereits bekannt ist. Dies ergibt die notwendige Einschränkung im laufenden Betrieb, definiert die zu betreuenden Fachgebiete und die Inhalte der Sammlung können vielfach aus dem Lehrbetrieb, in Form von Skripten, Arbeitsmaterialien, oder aus Literaturhinweisen, übernommen werden.

Beim konkreten Aufbau der Sammlung haben wir drei verschiedene Quellen für unseren Bestand definiert:

1. Sammlung verstreut vorliegender digitaler Dokumente
2. Erstellung neuer digitaler Dokumente
3. Digitalisierung konventioneller Dokumente

Bei der Erschließung der Inhalte griffen wir auf das intellektuelle Verfahren zurück, in Kombination mit den Methoden der Volltextindizierung. Wir sahen uns nach eingehender Diskussion nur so in der Lage, den verlangten Qualitätsanspruch zu erreichen. Die praktische Umsetzung dieses Konzeptes wird im nächsten Abschnitt an gegebener Stelle näher vorgestellt.

## 2 Umsetzung

Neben einer Konzeption für digitale Bibliotheken beschäftigte sich die Seminargruppe auch mit deren praktischer Umsetzung. Hierzu wurden verschiedene zu diesem Zweck verfügbare Software-Lösungen näher betrachtet und diskutiert. Neben der letztendliche Lösung für das Projekt wird hier noch ein in anderem Rahmen an der Fachhochschule eingesetztes Produkt vorgestellt.

### 2.1 Hyperwave Information Server

#### 2.1.1 Allgemeines

Der „Hyperwave Information Server“ (HIS)<sup>3</sup> ist eigentlich nicht für den Einsatz zur Realisierung einer digitalen Bibliothek gedacht. Im Prinzip entspricht seine Architektur der einer Groupware. Ziel ist es, verschiedenen Nutzern abgestuften Zugriff auf die Verzeichnisse zu ermöglichen. Hintergrund ist das Ziel, Datenbestände autark und dezentral pflegen zu lassen, was vor allem im Bereich mittlerer bis großer Intranets interessant ist.

Zu diesem Zweck verfügt der HIS über recht umfangreiche endnutzerorientierte Dialoge, mit denen eine Vielzahl der Operationen ausgeführt werden können. Diese Funktionalität ist zweiseitig durch JavaScript- und serverseitige Anwendungen realisiert. Im Kern umfasst dies das Heraufladen, Bearbeiten und Löschen von Dateien.

Die Speicherung des Datenbestandes erfolgt in einer objektorientierten Datenbank. Verknüpfungen innerhalb von Dokumenten werden dabei als Objekte gespeichert. Dies ermöglicht eine Kontrolle und z. B. auch das Entfernen toter Verknüpfungen. Eine Volltextindizierung der Dokumente erfolgt zeitgleich mit deren Einstellung, so dass diese sofort retrievelfähig sind.

Die Basisfunktionalität ist jedoch die eines HTTP-Servers, bei dem die Objekthierarchie auf eine Verzeichnisstruktur abgebildet wird. Das Gestalten der Oberfläche entspricht daher weitgehend der Arbeit an herkömmlichen web-basierten Serverlösungen. Per Voreinstellung wird die Objekthierarchie in einem an Microsoft Windows angelehnten Aussehen dargestellt.

#### 2.1.2 Umsetzung

Im Zuge der konkreten Umsetzung wurden einige Mängel am Konzept der Software erkennbar. Vor allem war dies die erhebliche Last auf der Client-Seite, die beim Einsatz anfiel. Die Anwendung mit einem aktuellen Netscape-Browser auf einer Pentium-Maschine bei einer Taktung von 90 Megahertz und 32 Megabytes Arbeitsspeicher wurde für die Beteiligten zur Tortur, da die Arbeit von ständigen Browser-Abstürzen begleitet wurde.

Die Vergabe von Metadaten und die weiteren Aktionen konnten zwar über Dialoge realisiert werden, die Gestaltung der Seiten zur Darstellung musste jedoch manuell erfolgen. Der Zugriff auf die zusammen mit dem Dokument repräsentierende Datenbank-Objekt gespeicherten Metadaten war über eine server-spezifische proprietäre Skriptsprache möglich, die allerdings im uns gesetzten Rahmen zu umständlich für den Einsatz war.

---

<sup>3</sup><http://www.hyperwave.de>

Die manuelle Erstellung von Einstiegseiten zu den Dokumenten war deshalb notwendig, um dem Zugriff notwendige Informationen zu diesen Dokumenten vorzuschalten. Der Zugriff über eine – zwar kommentierte – Verzeichnisstruktur erschien in diesem Zusammenhang nicht angemessen.

Ein weiteres Hindernis war die verhältnismässige schlechte Leistung des Servers, da die Anwendung sehr rechenintensiv ist. Dies war auch durch die Verwaltung der hinter dem Server stehenden Datenbank bedingt, die zudem einen recht großen Bedarf an Massenspeicher entwickelte. Hinzu kommt noch ein durch zahlreiche Grafikdateien und verschachtelte Frames stark erhöhtes Datenvolumen.

### 2.1.3 Schlussfolgerungen

Zusammenfassend kann gesagt werden, dass der HIS sich für groupware-basierte Arbeiten in Intranets auf Grund seiner Rechteverwaltung und des dezentralen Ansatzes anbietet, aber die server- als auch client-seitigen Anforderungen an die Systemlast ihren entsprechenden Preis haben. Dies hat zur Folge, dass zwar aktuelle, aber nicht sehr neue Systeme für einen Einsatz nicht in Frage kommen.

Im Content-Management-Umfeld sind in der Zwischenzeit vielfältige und leichtgewichtigere Lösungen verfügbar, die in vielen Fällen die gleiche Funktionalität bieten. Zudem ist für den alleinigen Einsatz zum Zwecke einer digitalen Bibliothek ein dedizierteres System wünschenswert.

Diese Gründe führten zu der Entscheidung gegen den HIS als Lösung für das hier vorgestellte Projekt.

## 2.2 Greenstone Digital Library Software

Um den Einsatz der GSDL-Software ist letztlich das hier vorgestellte Projekt entstanden. Daher wird diese Lösung hier ausführlicher dargestellt.

### 2.2.1 Allgemeines

Greenstone ist eine im Rahmen des *New Zealand Digital Library* Projektes entstandene und unter der GPL<sup>4</sup> stehende Software zur Erstellung digitaler Bibliotheken. Das Paket ist über die Seite der NZDL<sup>5</sup> erhältlich.

Besonders nennenswerten Einsatz findet Greenstone beim Aufbau der „*Humanity Development Library*“ (Humanity Libraries Project 1999 u. ö.), eines von mehreren Institutionen kooperativ erstellten und von der UNESCO geförderten Projektes.

Ein Bonus ist die Möglichkeit, Inhalte auf CD-ROM zu exportieren. Auf diese Weise erstellte CD-ROMs sind unter Microsoft Windows nutzbar und bieten dort die gleiche Oberfläche und Funktionalität wie im Netzzugriff.

Greenstone ist auf diversen Unix- und Microsoft-Windows-Systemen lauffähig und stellt relativ geringe Systemanforderungen. Der Testbetrieb in unserem Projekt erfolgte auf einer Maschine mit zwei Intel Pentium-Prozessoren bei einer Taktung von jeweils 90 Megahertz und 64 Megabytes Arbeitsspeicher. Die Festplatte hatte eine Speicherkapazität von 500 Megabytes. Als Betriebssystem kam die Debian 2.2 Distribution von GNU/Linux mit Erweiterungen auf den

---

<sup>4</sup> „GNU General Public License“ (s. Free Software Foundation 1991).

<sup>5</sup><http://nzdl.org>



Kernel 2.4 zum Einsatz. Da Greenstone nur als CGI-Backend realisiert ist, war der Einsatz eines HTTP-Servers von Nöten. In unserem Fall war dies der Apache in der Version 1.3.9. Im Test lief das System zwar merklich unter Last, jedoch immer noch mit zumutbaren Reaktionszeiten.

Die Software ist auf Mehrsprachigkeit ausgelegt und unterstützt zahlreiche Zeichensätze und Sprachanpassungen.

### **2.2.2 Software-Konzept**

Greenstone sieht eine Einteilung des Bestandes in verschiedene, von einander getrennte „Collections“, also Sammlungen, vor. Dies verlangt schon bei der Planung gezielte Überlegungen über eine Aufteilung des Bestandes. Die im Rahmen unseres Projektes angelegten Sammlungen unterscheiden sich sowohl nach inhaltlichen und/oder formalen Aspekten.

Die Nutzung und der Zugang erfolgt nach dem Client-Server-Prinzip über das HTTP-Protokoll mit den üblichen Web-Browsern. Die Gestaltung der Benutzeroberfläche erfolgt automatisch nach Vorlagen und der Struktur der eingestellten Dokumente. Hier sind sogenannte „Classifier“ vorgesehen, die Index-Listen aus angegebenen Metadaten-Feldern erstellen, in denen der Benutzer wie in einem Bibliothekskatalog blättern („browsen“) kann.

Kern der Software ist das Indizierungs- und Retrievalprogramm „mg“, das bei Witten u. a. (1999) beschrieben wird. Damit steht ein leistungsfähiger Kern zur Verfügung, der, wie wir noch zeigen werden, auch komplexe Suchanfragen in großen Datenbeständen effizient verarbeiten kann. Indizes können hiermit aus beliebig wählbaren Dokumentelementen und Metadaten gebildet werden.

Das Einstellen der Dokumente wird über sogenannte „Plug-Ins“ realisiert, die einen Zugang zu diversen Datei- und Speicherformaten erlauben und bei Bedarf auch eigenständig angepasst oder komplett neu erstellt werden können.

Greenstone ist auf den Einsatz von Metadaten ausgelegt. Zwar wird Dublin Core bevorzugt und unterstützt, prinzipiell lässt sich aber jede Form von Metadatenformaten im System verwenden.

Zur Pflege der Sammlungen wird zwar eine grafische Oberfläche angeboten, die allerdings nur für Endnutzer gedacht ist, die ohne Mühe eigenen Dokumente zu der heimischen Installation hinzufügen wollen. Für den professionellen Einsatz muss auf Kommandozeilen-Befehle zurück gegriffen werden, die nicht unbedingt benutzerfreundlich, dafür aber umso effektiver sind.

Eingestellte Dokumente werden komprimiert und, abhängig von der Konfiguration, indiziert gespeichert. Das Konzept der Komprimierung ermöglicht bei den angewandten Verfahren einen mühelosen Umgang auch mit sehr großen Datenmengen.

### **2.2.3 Beispielanwendungen**

Für den Aufbau der Beispielanwendungen wurde die Greenstone-Software in der Version 2.35 eingesetzt. Gegen Ende der Projektphase 2 wurde die Version 2.36 verfügbar und testweise parallel auf einem Heim-PC betrieben.

Zur Veranschaulichung der Möglichkeiten wurden drei Sammlungen mit verschiedenen Anforderungen aufgebaut, die wir im Folgenden darstellen wollen.

**„GNU/Linux Dokumentation“** Die Sammlung mit technischer Dokumentation aus dem Bereich des GNU/Linux-Betriebssystems war weniger auf die qualitative Erschließung ausgelegt, als auf die schnelle Zur-Verfügung-Stellung von relevanten Dokumenten. Der größte Teil des Bestandes wurde aus dem *Linux Documentation Project*<sup>6</sup> übernommen, der Rest stammt aus einschlägigen Quellen.

Der Bedarf an diesen Dokumenten legt keine tiefgehende Erfassung nahe, zudem die Inhalte eine geringe Halbwertszeit aufweisen. Darüber hinaus wurde auch die typische Herangehensweise der (potentiellen) Benutzer berücksichtigt, als auch die besondere Gestaltung des Materials: der Inhalt der Dokumente ist immer sehr spezifisch auf ein Thema hin ausgelegt, der Titel ist entsprechend eindeutig und der überwiegende Teil enthält aussagekräftige Abstracts. Die Herangehensweise des Nutzers ist fast ausschließlich auf den Titel fixiert. So war die Erfassung von Titel, Autor, Datum und Abstract ausreichend, um die Relevanz eines Dokumentes in der Aufnahme erkennen zu lassen, zudem alle Daten direkt aus dem Dokument per „Copy & Paste“ übernommen werden konnten. Nicht einmal die Herkunft der Dokumente wurde erfasst, da diese in der Regel auf den ersten Seiten extra genannt ist, z. B. in der Form

The newest version of this document can be found at <URL>.

Die inhaltliche Erschließung erfolgte rein auf Basis der Volltextindizierung. Ein Register wurde für den Titel angelegt.

Das Quellformat der Dokumente ist recht homogen, da Dokumentationen in diesem Bereich fast ausschließlich in einer SGML-Anwendung erstellt werden und anschließend als PDF-, HTML-, Text- oder PostScript-Dateien zur Verfügung gestellt werden. Im vorliegenden Fall besteht die Sammlung aus PDF- und PostScript-Dateien.

Bei der Einstellung der Dokumente extrahierten die Plugins den Text aus den Dateien und speicherten diesen – komprimiert – im HTML-Format<sup>7</sup>. Daher existieren zwei Versionen eines Dokumentes innerhalb der Sammlung: einmal das Original-Dokument und der Greenstone-Text. Letzterer hat den Vorteil, das Suchbegriffe hervorgehoben werden können. Allerdings eignet er sich auf Grund seiner schlechten Formatierung zu wenig mehr, als dem schnellen Prüfen des Textes, das direkt im Browser und ohne Plugin erfolgen kann. Aufgrund der letztlich doch relativ schlechten Verwertbarkeit für den Nutzer ist daran zu denken, den Zugriff auf diesen Text zu entfernen. Dies vermeidet zudem noch Irritationen.

Nach einem ersten Probelauf betrug die Größe aller Dokumente zusammen 25 Megabytes, der komprimierte Text und die Indizes nochmals 16 Megabytes. Eine Suche in diesem Bestand nach den (noch nicht gestoppten) Stoppwörtern „and or not but a an the one ein eine einer der die das“ lieferte nach einer Reaktionszeit von nur drei Sekunden die Ergebnisse. Ein von der Leistung her nach unserem Dafürhalten respektables Ergebnis, besonders hinsichtlich der Leistungsgrenze der Servermaschine.

**„Digitale Bibliotheken“** Diese Sammlung bildet den Kern des Projektes. Sie stellt den Aufbau eines qualitativ hochwertig erschlossenen Dokumentenbe-

---

<sup>6</sup><http://www.linuxdoc.org>

<sup>7</sup>Korrektur: in einem greenstone-eigenen Auszeichnungs-Format

standes dar. Das thematisch definierte Sammlungsgebiet umfasst Dokumente zum Komplex der digitalen Bibliotheken im weiteren Sinne.

Die Wahl des Formates war lediglich durch die technischen Möglichkeiten der Greenstone-Software eingeschränkt. Gesammelt werden PDF-, PostScript- und HTML-Dateien. Das Einspielen von Microsoft-Word-Dateien war allerdings in Abhängigkeit von der Format-Version mit Problemen verbunden. Dies und die Zuordnung zu den proprietären Formaten war Grund genug, es nicht zu unterstützen. Nicht lösbare Probleme brachten an Frames gebundene HTML-Dateien mit sich, da diese nicht als Ganzes und ohne Informationsverlust importiert werden konnten.

Erschlossen werden die Dokumente in enger Anlehnung an Dublin Core (1999):

**Creator** Verfasser des Artikels in invertierter Form, allerdings nach Vorlage. Vorerst werden noch alle Verfasser übernommen.

**Title** Titel nach Vorlage. Hierbei werden Einzeldateien auch als Einzelwerke angesehen. Bei Kapiteln, die zu einem übergeordneten Werk gehören wird der Titel dieses Werkes als Gesamttitel erfasst und die Ordnungsnummer der Kapitel als Zählung.

**Date1** Datum der Erstellung des Dokumentes. Wenn kein Datum angegeben war, wurde das letzte greifbare Datum verwendet, als letzte Lösung der „*Last-Modified*“-Header aus dem HTTP-Dialog.

**Date2** Datum der Einstellung in die Sammlung; Zugangsdatum.

**Subject** Aus der SWD vergebene Schlagwörter, im Allgemeinen jedoch ohne Berücksichtigung der RSWK. Statt dessen wurde eher auf eine umfassende Repräsentation des Inhaltes Wert gelegt.

**Description** Kurzreferat. nach Möglichkeit aus dem Dokument übernommen, ansonsten eigens verfasst.

**Class** Systemstelle nach eigenem Schema.

**Source** Quell- oder Zitieradresse des Dokuments.

**Rights** Urheberrechtsvermerk.

Bei der Suche nach einem geeigneten Klassifikationsschema deckten alle von uns untersuchten Lösungen das Gebiet der digitalen Bibliotheken nicht in der wünschenswerten oder notwendigen Tiefe ab. Aus diesem Grund entwarf einer der Seminarteilnehmer eine speziell auf dieses Fachgebiet ausgerichtete Dezimalklassifikation, die über dieses Projekt hinaus für den produktiven Einsatz entwickelt werden soll.

Der Zugang zu den Dokumenten erfolgt entweder über Register oder über eine Suchfunktion. Zur Suche wurden Indizes aus den Metadaten und aus dem Volltext angelegt, die jeweils nur getrennt abfragbar sind. Die Register wurden in Anlehnung an konventionelle Kataloge erstellt:

**Autor** alphabetische Liste aus *Creator*

**Titel** alphabetische Liste aus *Title*

## **Sachgruppen** hierarchische Liste aus *Class*

Diese verschiedenen und sich ergänzenden Zugänge zu der Sammlung bieten einerseits aus den konventionellen Bibliotheken bekannte Verfahren, um Dokumente zu finden, andererseits schöpfen die Indizierungsverfahren das Potential der digitalen Bibliothek voll aus.

Angestrebt ist, das Material auf lange Zeit zu archivieren. Das heißt, neue Versionen eines Dokumentes werden hinzugefügt, die alten aber nicht gelöscht.

**„Computer-Bibliografie“** Die „Computer-Bibliografie“ ist nicht wie die anderen Sammlungen zur Speicherung von Dokumenten gedacht, sondern zum bibliografischen Nachweis von für den IuK-Bereich relevanter EDV-Literatur. Erfasst wurden die Angaben im „GNU Bib/Refer“-Format (vgl. Jacobsen 1995), das für unsere Zwecke um wenige Felder erweitert wurde.

Bei der Erfassung der Literatur stand die Regelwerkskonformität im Hintergrund. Auch hier sollten die Daten schnell und einfach in einem zitierfähigen Format erfasst werden.

Der Betrieb parallel zu einem Bibliothekskatalog wäre dann fraglich, wenn beide Systeme den gleichen Bestand aufwiesen. Die Idee war jedoch, in der Computer-Bibliografie jedwede für unseren Bereich relevanten Titelangaben, möglichst mit Kurzreferat, aufzunehmen. Das hieße auch ohne Autopsie, z. B. nach Buchhandelskatalogen, Literaturverzeichnissen, Prospekten, etc. Da die so angegebene Literatur vielfach nicht in der Bibliothek vorhanden wäre, und dies auch oft nicht sein muss, wurden die Angaben um eine Verknüpfung zu Lehmanns Fachbuchhandlung<sup>8</sup> ergänzt, so dass der entsprechende Titel recht schnell privat bestellt werden kann.

Im Prinzip kann auf diese Weise bei kontinuierlicher Pflege eine aktuelle Bibliografie der IuK-relevanten EDV-Literatur aufgebaut werden.

**Export auf CD-ROM** Eher im Hintergrund während des Projektes stand die Möglichkeit, Sammlungen auf CD-ROM zu exportieren. Bei einem solchen Vorgehen legt die Greenstone-Software ein Paket an, das den gesamten Datenbestand einer Sammlung enthält, sowie alle zur Installation und Nutzung unter Microsoft Windows notwendigen Anwendungen und Programmbibliotheken. Nach der Installation ist über den Systemmanager eine Programmgruppe verfügbar, über welche die auf die Sammlung zugegriffen werden kann. Dabei wird lokal ein Server gestartet, der die native Web-Umgebung emuliert. Die Sammlung selbst kann dabei auf die Festplatte kopiert oder von CD-ROM geladen werden.

Diese Möglichkeit verdiente eine genauere Betrachtung, da das hierin steckende Potential gewinnbringend und sinnvoll genutzt werden kann, wie dies auch schon an anderer Stelle getan wird (vgl. UNESCO 2000)

## **2.3 Resumée**

In unserem Fall ist zu sehen, dass es keine perfekte Software-Lösung gibt, und dass an der einen und anderen Stelle immer wieder Zugeständnisse an die Technik notwendig sind. Jedoch ist es möglich, bei genauerer Betrachtung der Sachlage Kriterien zu definieren, nach denen der Einsatz einer Lösung bewertet werden

---

<sup>8</sup><http://www.lob.de/?flag=hbi>

kann. In unserem Fall fiel die Entscheidung zu Gunsten der Greestone-Software, wie wohl diese nicht der Stein der Weisen ist, jedoch unseren Anforderungen genügte und sich sehr einfach in den Arbeitsfluss integrierte.

Als unverzichtbar stellte sich die Ausrichtung auf freie Software heraus. Hier hat sich gezeigt, dass auftretende Probleme durch Anpassungen des Quellcodes rasch lösbar sind, wie auch nicht oder schlecht dokumentierte Eigenschaften der Software durch Studium dieses Codes nachvollziehbar werden.



### 3 Ausblick

Der Einsatz spezialisierter, oder als solcher bezeichneter, Software darf nicht alleine deswegen erfolgen, weil man es machen kann. Derartige Ansätze sind leider oft zu beobachten. In der allgemeinen Manie um alles, was mit „*Medien*“ zu tun hat, geht leider das Verständnis von einem Medium als „vermittelndem Element“ unter, und damit auch die Bedeutung der Inhalte, die auf diesem Wege vermittelt werden sollen.

Beim Aufbau der digitalen Bibliothek hat sich gezeigt, dass Erfahrungen aus dem Aufbau und der Pflege konventioneller Bestände hier als Stärke zu Tage treten und sich die Herangehensweise der Bibliothekare und Dokumentare durchaus bewährt hat. In diesem Sinne hat sich auch deren klassisches Berufsbild verändert. Mehr und mehr informationstechnische Kenntnisse sind notwendig, um Lösungen effektiv bewerten und einsetzen zu können.

Nach Abschluss des Großteils dieses Berichts veröffentlichte der Wissenschaftsrat eine „Empfehlung zur digitalen Informationsversorgung“ (Wissenschaftsrat 2001) und auch die schon länger zugängliche SteFi-Studie (Klatt u. a. 2001) macht die Bedeutung digitaler wissenschaftlicher Information und dem professionellen Zugang dazu deutlich. In digitalen Bibliotheken steckt ein in noch vielen Bereichen ungenutztes aber vorhandenes Potential, das sich zur qualifizierten Informationsvermittlung gewinnbringend nutzen ließe. Eine sehr gute Darstellung dieses Themas findet sich bei Endres und Fellner (2000).

Im Umfeld von oft zweifelhaften Ansätzen und dem verloren gegangenen Element der „Bibliothek“ sind vor dem Reiz des technischen Machbaren überzeugende Beispiele leider recht selten. Vor diesem Hintergrund ist unser Modell nicht als Argument, sondern als Beispiel gedacht, dessen Funktionalität sich im nun kommenden Produktivbetrieb zu beweisen hat.

Zum Zeitpunkt dieses Berichtes ist die Aufbauphase unserer digitalen Bibliothek noch nicht abgeschlossen. Der Betrieb des hier entwickelten Angebotes soll von der Bibliothek und dem Informationstechnischen Service weitergetragen werden und bis November 2001 öffentlich verfügbar sein. Informationen über den Stand des Projektes sind mit der Veröffentlichung dieses Berichtes unter der Adresse <http://diana.iuk.hdm-stuttgart.de/digbib> verfügbar.





## Literaturverzeichnis

- Adobe 2000** ADOBE SYSTEMS INC.: *PostScript*. 2000. – URL <http://partners.adobe.com/asn/developer/technotes/postscript.html>. – Zugriffsdatum: 2001-03-27
- Adobe 2001** ADOBE SYSTEMS INC.: *Technical notes : Acrobat/PDF*. 2001. – URL <http://partners.adobe.com/asn/developer/technotes/acrobatpdf.html>. – Zugriffsdatum: 2001-03-27
- Berners-Lee 1999** BERNERS-LEE, Tim: *Der Web-Report*. München : Econ, 1999. – 332 S. – ISBN 3-430-11468-3
- Biersch 2000** BIRSCH, Stefanie: *E-Journals im Internet und deren Einfluß auf die Funktion der Bibliothek der Knoll AG Ludwigshafen : unter besonderer Berücksichtigung der Freeprint- und Preprint-Archive*, FH Stuttgart – HBI, Diplomarbeit, 2000. – 92 S. S
- Born 2000** BORN, Günter: *Dateiformate : die Referenz*. 1. Aufl. Bonn : Galileo Computing, 2000. – 1127 S. – ISBN 3-934358-83-7
- DANTE 1994 u. ö.** : *Deutschsprachige Anwendervereinigung T<sub>E</sub>X e. V.* 1994 u. ö.. – URL <http://www.dante.de>. – Zugriffsdatum: 2001-02-20
- DBI 1996 u. ö.** EXPERTENGRUPPE RAK: *Regeln für die alphabetische Katalogisierung von Nichtbuchmaterialien : RAK-NBM*. Berlin : DBI, 1996 u. ö.. – Loseblatt-Ausg.
- DIN e. V. u. a. 2001** DIN E. V. u. a.: *Anleitung zur Nutzung von Dublin Core in Europa*. Berlin : Beuth, 2001. – Norm-Nr.: DIN CWA 13988:2000
- Dublin Core 1999** DUBLIN CORE METADATE INITIATIVE: *Dublin Core element set, Version 1.1*. 1999. – URL <http://www.dublincore.org/documents/dces/>. – Zugriffsdatum: 2001-04-10
- Endres und Fellner 2000** ENDRES, Albert ; FELLNER, Dieter W.: *Digitale Bibliotheken : Informatik-Lösungen für globale Wissensmärkte*. 1. Aufl. Heidelberg : dpunkt-Verl., 2000. – 494 S. – ISBN 3-932588-77-0
- Free Software Foundation 1991** FREE SOFTWARE FOUNDATION: *GNU General Public License*. 1991. – URL <http://www.fsf.org/gpl/>. – Zugriffsdatum: 2001-02-20
- Humanity Libraries Project 1999 u. ö.** HUMANITY LIBRARIES PROJECT: *Humanity Development Library*. 1999 u. ö.. – URL <http://www.humaninfo.org/>. – Zugriffsdatum: 2001-06-13
- ISO 1986** INTERNATIONAL ORGANISATION FOR STANDARDIZATION: *Information processing - Text and office systems - Standard Generalized Markup Language (SGML)*. 1. ed. Geneva : ISO, 1986. – XI, 155 S. S. – Norm-Nr.: ISO 8879-1986
- Jacobsen 1995** JACOBSEN, Dana: *Bibliography formats*. 1995. – URL <http://www.ecst.csuchico.edu/~jacobsd/bib/formats>. – Zugriffsdatum: 2001-06-25

- Klatt u. a. 2001** KLATT, Rüdiger u. a.: Nutzung elektronischer wissenschaftlicher Information in der Hochschulausbildung / BMB+F. URL <http://www.stefi.de>. – Zugriffsdatum: 2001-07-15, 2001. – Forschungsbericht
- Nelson u. a. 1960 u. ö.** NELSON, Ted u. a.: *The Xanadu Project*. 1960 u. ö.. – URL <http://www.xanadu.net>. – Zugriffsdatum: 2000-11-29
- UNESCO 2000** UNESCO: Digital Libraries. In: *UNESCO Sources* June 2000 (2000), Nr. 124, S. 12. – ISSN 1014-6989
- W3C 1995 u. ö.** W3C: *Hypertext markup language : home page*. 1995 u. ö.. – URL <http://www.w3.org/MarkUp/>. – Zugriffsdatum: 2000-12-07
- W3C 2000** W3C: *Extensible markup language (XML) 1.0*. 2000. – URL <http://www.w3.org/TR/2000/REC-xml-20001006>. – Zugriffsdatum: 2000-12-07
- Wissenschaftsrat 2001** WISSENSCHAFTSRAT: *Empfehlung zur digitalen Informationsversorgung durch Hochschulbibliotheken*. 2001. – URL <http://www.wissenschaftsrat.de/texte/4935-01.pdf>. – Zugriffsdatum: 2001-07-15
- Witten u. a. 1999** WITTEN, Ian H. ; MOFFAT, Alistair ; BELL, James: *Managing gigabytes : compressing and indexing documents and images*. 2nd ed. San Francisco : Morgan Kaufmann, 1999. – XXXI, 519 S. S. – ISBN 1-55860-570-3
- Zimmer 2000** ZIMMER, Dieter E.: *Die Bibliothek der Zukunft*. 3. Aufl. Hamburg : Hoffmann und Campe, 2000. – 331 S. – ISBN 3-455-10421-5

## Kolophon

Dieser Text wurde gesetzt mit  $\text{\LaTeX} 2_{\epsilon}$  aus der 10 Punkt Computer-Modern-Schrift. Für den Satz wurden die `ec`-Schriften von Jörg Knappen verwendet. Die PDF-Version dieses Dokumentes verwendet aus Mangel an Type-1-Schriften noch die Standardvariante der Computer-Modern-Fonts. Gegenüber der PostScript- und damit der Druckfassung unterscheidet sich diese in typografischen Details wie Unterschneidungen und Wortumbrüchen.

Das Literaturverzeichnis wurde nach DIN 1505 mit  $\text{\BibTeX}$  und dem `dinat`-Paket von Helge Baumann (`helge.baumann@gmx.de`) erstellt.

Zur vollständigen Auflösung aller Referenzen benötigte  $\text{\LaTeX}$  insgesamt vier Läufe.

Der Satz erfolgte auf einer unter GNU/Linux betriebenen Pentium-Maschine mit 233 Megahertz Taktung unter Verwendung der  $\text{teTeX}$ -Distribution in der Version 1.0.6.

Alle Eingabedateien brachten eine Summe von 81 Kilobytes auf. Die Größe der Ausgabedateien betrug im DVI-Format 67 Kilobytes, im PS-Format 251 Kilobytes und im PDF-Format 223 Kilobytes.

